

Db 421 CTTTGGCCCTTTGTGTTGGCGATACGATCAGGCGTAATAAGCCTTTCAAGAATGTGTCATC 480

Qy 552 tgatactgagacatttgttataccggatttcccgcagtgagctgaagtttgtgaggactca 611
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 481 GGATACTGAAACTTTTGTGTACCGGATTGCCACATGAAATTAGGCTAACTAGAACACA 540

Qy 612 agtggctccgtttcagcttgcggaacggagaatggattctcaaagttgatgaaacagat 671
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 541 GTTGTCTCCGTTTGAGCAATCGGATGAAGAGACGGGTATGGCTCCCATGATTAAAGCTGT 600

Qy 672 gacggagtctgttggtagaagctacggtgttggttaacagtttttatgagctcgagtc 731
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 601 GAGGGAATCGGATGCGAAGAGCTATGGAGTTATATTCAATAGCTTTTATGAGCTTGAATC 660

Qy 732 gacttatgtggattattacagagaggttttggtagaaagtcttggaatatagggcctct 791
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 661 AGATTATGTTGAACATTACACTAAGTTGTAGGTAGAAAAAATTGGGCTATTGGTCCGCT 720

Qy 792 gttgttatccaacaatggcaatgaggaaaaagtacaaaaggggaaaggaatctgcgattgg 851
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 721 TTCGCTGTGCAATAGGGATATTGAAGATAAAGCGGAAAGAGGGAGGAAATCATCTATCGA 780

Qy 852 cgaacacgaatgcttggttggtgaattccaagaagcagaattcggttggttacgtttg 911
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 781 TGAACACGCGTGCTTGAAATGGCTTGATTCTGAAGAAATCAAGTTCCATTGTTTATGTTTG 840

Qy 912 ttttggaagtatggcgacttttactccagcgcagttgcgcgaaaactgcgattggactcga 971
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 841 TTTTGGAAGTACAGCAGATTTCACACAGCAGATGCAAGAACTTGCTATGGGGCTAGA 900

Qy 972 ggaatcaggccaagagttcatttggttagttaaaaaaggccaaaaacgaagaagaaggaaa 1031
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 901 AGCCTCTGGACAAGATTTTCAATTTGGGTTATCA-----GAAC 936

Qy 1032 aggaaaaagaagaatggctgccagaaaattttgaggaaagagtgaagatatagggttgat 1091
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 937 AGGGAATGAAGATTGGCTCCCAGAAGGATTCGAGGAAAGAACAAAAGAAAAGGTTTAAT 996

Qy 1092 cataagaggatgggcgccgcaattgttgatactcgatcatcctgcggtaggagctttcgt 1151
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 997 CATAAGAGGATGGGCACCCCAAGTGTGATTCTTGATCACGAAGCTATTGGAGCTTTTGT 1056

Qy 1152 gacgcattgtggatggaattcgacggttgaaggaatatgcgccggtgtgcctatggtgac 1211
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1057 TACTCATTGTGGATGGAACCTGACACTGGAAGGAATATCAGCAGGGGTACCAATGGTGAC 1116

Qy 1212 ttggccagttttcgcagagcagtttttcaatgagaagtttgtgacagaggttttggggac 1271
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1117 ATGGCCAGTATTTGCGGAACAGTTTTTCAATGAGAAGTTGGTGACTGAGGTAATGAGAAG 1176

Qy 1272 cggtgtttcggttggaataagaagtggctaagggcagcaagtgaaggtgtgtcgaggga 1331
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1177 TGGAGCTGGTGTTGGTTCTAAGCAATGGAAGAGAACAGCTAGTGAAGGAGTGAAAAGAGA 1236

Qy 1332 ggcagtgacgaacgcggtgcagcgtgttatggtgggagaaaatgcgtcggagatgagaaa 1391
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1237 AGCAATAGCAAAGGCGATAAAGAGAGTAATGGCGAGTGAAGAAACAGAGGGATTGAGAAG 1296

Qy 1392 gcgagcgaagtattataaggaaatggcgaggcgggcggttgaggaaggcgggttcgtctta 1451
 ||||| || | ||| || ||||| || || ||||| ||||| ||||| |||||
 Db 1297 CAGAGCAAAAGAGTACAAAGAAATGGCAAGAGAAGCTATTGAAGAAGGAGGATCATCTTA 1356

 Qy 1452 taatggtttgaatgagatgatagaggatttgagtgtgtaccgtgctccagaaaaacaaga 1511
 ||||| || | ||||| || || ||||| || || ||||| || || |||||
 Db 1357 CAATGGATGGGCTACTTTGATACAAGACATAACTTCATATCGTTAACTAGTTGATGCAAA 1416

 Qy 1512 cttaaactagattccttatagatgacttctagtgtgacaattgtaattttttgccttttat 1571
 || || || || || || || || || || || || || || || || || || || || ||
 Db 1417 AAAAGAAAAACATGTGTGTTTCTATATTCTGTCTTCTGTTTTGCTGATTTGATCATATT 1476

 Qy 1572 tcaagtttcctcattagtgttga 1594
 | ||| |||| | |||
 Db 1477 ACGTACTTCTTCATGATAATTAA 1499

RESULT 7

AAW47172

ID AAW47172 standard; Protein; 470 AA.

XX

AC AAW47172;

XX

DT 08-JUN-1998 (first entry)

XX

DE Glucosyl transferase (GTase) protein encoded by TWI1 gene.

XX

KW Glucosyl transferase; GTase; TWI1; tomato; signalling pathway;

KW salicylic acid; jasmonic acid; ethylene; wound inducible gene;

KW plant defence protein; plant response; tobacco; rice.

XX

OS Lycopersicon sp.

XX

PN WO9745546-A1.

XX

PD 04-DEC-1997.

XX

PF 30-MAY-1997; 97WO-GB01473.

XX

PR 31-MAY-1996; 96GB-0011420.

XX

PA (UYYO-) UNIV YORK.

XX

PI Bowles DJ, Calvert CM, O'Donnell PJ, Roberts MR;

XX

DR WPI; 1998-032653/03.

DR N-PSDB; AAV17054.

XX

PT Tomato wound inducible (TWI1) gene encoding glucosyl transferase -

PT useful to develop products that alter signalling pathways in plants

PT by altering of salicylic acid, jasmonic acid or ethylene

XX

PS Claim 2; Fig 3; 52pp; English.

XX

CC This is a glucosyl transferase (GTase) protein encoded by a wound
 CC inducible gene (TWI1) isolated from wounded tomatoes. The TWI1 gene
 CC encodes this GTase from amino acid position 5. The TWI1 gene can be
 CC used to identify homologue GTase encoding genes isolated from tobacco
 CC and rice. A microbial host can be transfected or transformed with a
 CC vector containing the GTase encoding nucleic acids. The products can be
 CC used to interfere with GTase and therefore alter signalling pathways in
 CC plants, specifically tobacco, rice or tomato plants by altering levels
 CC of salicylic acid, jasmonic acid or ethylene. This can induce the
 CC production of plant defence proteins such as pathogenesis-related (PR)
 CC and proteinase inhibitor (PIN) proteins which regulate plant development
 CC (plant growth, reproduction and senescence) and improve plant response to
 CC pathogens.

XX

SQ Sequence 470 AA;

Query Match 59.2%; Score 1490.5; DB 19; Length 470;

Best Local Similarity 57.3%; Pred. No. 1.2e-144;

Matches 274; Conservative 84; Mismatches 107; Indels 13; Gaps 3;

Qy 1 MGKHLHIALFPVMAHGHI PMLDMAKLFTSRGIQTTIIST----LAFADPINKARDSGLDI 56
 ||:|| || | |||| |||| : ||:|| ||:|| | : | : : ||:||
 Db 1 mgelhffffpddagghmiptldmanvvacrgvkatiiittplnesvfskaiernkhlgiei 60

Qy 57 GLSILKFPPEGSGIPDHMVSLDLV-TEDWLPKFVESLVLQEPVEKLI EELKLDCLVSDM 115
 : :||| : : ||: |||| ::| || |::: :::: |:|| | : |||||
 Db 61 dirllkfpakendlpedcerldlvpsddklpnflkaaammkdefeeligeocrpdclvsdm 120

Qy 116 FLPWTVDCAAKFGIPRLVFHGTSNFALCASEQMKLHKPYKNVTSDETETFVIPDFPHELKF 175
 |||| | |||| |||:||||| |||| : :: :||:|||:|||||:||| |||::
 Db 121 flpwttlsaakfsiprivfhgtsyfalcvgdtirrnpkfknvssdtetfvvpdlpheirl 180

Qy 176 VRTQVAPFQLAETENGFSKLMQMTE SVGRSYGVVNSFYELESTYVDYYREVLGRKSWN 235
 |||::||: :: | | : ::| : || :|||: ||||| ||::| :|:|||:
 Db 181 trtqlspfeqsdeetgmapmikavresdaksygvifnsfyelesdyvehytkvvgrknwa 240

Qy 236 IGPLLLSNNNGNEEKVQRGKESAIGEHECLAWLNSKKQNSVYVCFGSMATFTPAQLRETA 295
 |||| | | :| :||::|| | | || ||:|| :|:||||| | || ||::| |
 Db 241 igplslcnrdiedkaergrkssidehaclkwldskksssiyvvcfgstadfttaqmqla 300

Qy 296 IGLEESGQEFIWVVKKAKNEEEGKGKEEWLPENFEERVKDRGLIIRGWAPQLLILDHPAV 355
 :||| |||:||||: : | |:||| |||| |:||||| |||| :||| |:
 Db 301 mgleasgqdfiwvir-----tgnedwlpegfeertkekgliirgwapqsvildheai 352

Qy 356 GAFVTHCGWNSTLEGICAGVPMVTPVFAEQFFNEKFVTEVLGTGVSVGNKKWLRAASEG 415
 ||||| ||||| ||||| ||||| ||||| ||||: :| ||:|:| | ||||
 Db 353 gafvthcgwnstlegisagvpmvtpvfaeqffneklvtevmrsgagvgskqwktaseg 412

Qy 416 VSREAVTNAVQRMVGENASEMRKRAKYYKEMARRAVEEGSSYNGLNEMIEDLSVYR 473
 | |||: |::|| | | || | |||| |:||||| :|:||: ||
 Db 413 vkreaiaikaikrmaseetegfrsrakeykemareaieeggssyngwatliqditsyr 470

A68001
 LOCUS A68001 1624 bp DNA PAT 05-MAY-1999
 DEFINITION Sequence 1 from Patent WO9745546.
 ACCESSION A68001
 VERSION A68001.1 GI:4756806
 KEYWORDS .
 SOURCE unidentified.
 ORGANISM unidentified
 unclassified.
 REFERENCE 1 (bases 1 to 1624)
 AUTHORS Bowles,D.J., O'donnell,P.J., Roberts,M.R. and Calvert,C.M.
 TITLE USE OF A NOVEL GLUCOSYL TRANSFERASE
 JOURNAL Patent: WO 9745546-A 1 04-DEC-1997;
 UNIV YORK (GB)
 FEATURES Location/Qualifiers
 source 1. .1624
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 545 a 252 c 372 g 455 t
 ORIGIN

Query Match 31.3%; Score 547.8; DB 9; Length 1624;
 Best Local Similarity 61.9%; Pred. No. 4.5e-119;
 Matches 942; Conservative 0; Mismatches 542; Indels 39; Gaps 3;

Qy 87 tcacattgccttatttccagttatggctcatggtcacatgatcccaatgttggacatggc 146
 ||| || ||| ||||| || ||||| ||||| ||||| ||| | |||||
 Db 1 TCATTTTTTCTTCTTTCCCGATGATGCTCAAGGTCATATGATACCTACACTTGACATGGC 60
 Qy 147 caagctctttacctcaagaggcatacaaaacaacatcatttcgactctcgcc----- 198
 || | | | | || | | | | | | | | | | | |
 Db 61 GAACGTTGTCGCTTGTCTGGTGTAAAGCCACTATAATCACAAACACCTCTCAATGAATC 120
 Qy 199 ----ttcgctgatccgataaacaagctcgtgattcgggcctcgatattggactaagcat 254
 ||| || | | || | | | ||| || | | |||| | | |
 Db 121 TGTTTTCTCTAAAGCTATTGAGAGAAACAAGCATTTAGGTATTGAAATTGATATTCGTTT 180
 Qy 255 cctcaaatccaccagaaggatcaggaataccagatcacatgggtgagccttgatctagt 314
 || ||||| || | | | | | | | | | | | | ||||| ||
 Db 181 ACTAAATTCAGCTAAGGAGAATGATTTGCCTGAAGATTGTGAGCGTCTTGATCTTGT 240
 Qy 315 ---tactgaagattggctcccaaagtttggtgagtcattagtcttattacaagagccagt 371
 | |||| || ||||| || | | | | | | | | |||| | |
 Db 241 ACCTTCTGATGACAACTCCCAAACCTTCTTAAAGCTGCGGCTATGATGAAAGATGAATT 300
 Qy 372 tgagaagcttatcgaagaactaaagctcgactgtctcggttccgacatgttcttgcccttg 431
 |||| ||||| || |||| | || |||| ||||| || |||| | | ||
 Db 301 TGAGGAGCTTATTGGAGAATGTCGCCCTGATTGTCTTGTCTTGATATGTTCTTCCATG 360
 Qy 432 gacagtcgattgtgcggttaagttcggtattccgaggttggttttccacggaacgagcaa 491
 ||| ||| |||| || || || | || || || | || |||| ||||| || |
 Db 361 GACTACTGATAGTCAGCCAAATTTAGCATACCAAGAATTGTATTCCATGGAAGTAGTTA 420
 Qy 492 ctttgcgttggtgtgcttcggagcaaatgaagcttcacaagccttataagaatgtaacttc 551
 ||||| | |||| | || | || | | ||||| ||||| || ||
 Db 421 CTTTGCCCTTTGTGTTGGCGATACGATCAGGCGTAATAAGCCTTTCAAGAATGTGTCATC 480

Qy	552	tgatactgagacatttgtttataccggatttcccgcgatgagctgaagtttgtgaggactca	611
Db	481	GGATACTGAAACTTTTGTGTACCGGATTTGCCACATGAAATTAGGCTAACTAGAACACA	540
Qy	612	agtggctccgttttcagcttgcggaaacggagaatggattctcaaagttgatgaaacagat	671
Db	541	GTTGTCTCCGTTTGAGCAATCGGATGAAGAGACGGGTATGGCTCCCATGATTAAAGCTGT	600
Qy	672	gacggagtctgttggtagaagctacggtgttgtggttaacagtttttatgagctcgagtc	731
Db	601	GAGGGAATCGGATGCGAAGAGCTATGGAGTTATATTCAATAGCTTTTATGAGCTTGAATC	660
Qy	732	gacttatgtggattattacagagaggttttgggtagaaagtcttggaatatagggcctct	791
Db	661	AGATTATGTTGAACATTACACTAAGGTTGTAGGTAGAAAAAATTGGGCTATTGGTCCGCT	720
Qy	792	gttgttatccaacaatggcaatgaggaaaaagtacaaaggggaaaggaatctgcgattgg	851
Db	721	TTCGCTGTGCAATAGGGATATTGAAGATAAAGCGGAAAGAGGGAGGAAATCATCTATCGA	780
Qy	852	cgaacacgaatgcttggccttgggtgaattccaagaagcagaattcggttgtttacgtttg	911
Db	781	TGAACACGCGTGCTTGAAATGGCTTGATTCTGAAGAAATCAAGTTCCATTGTTTATGTTTG	840
Qy	912	ttttggaagtatggcgacttttactccagcgcagttgcgcgaaactgcgattggactcga	971
Db	841	TTTTGGAAGTACAGCAGATTTCACTACAGCACAGATGCAAGAACTTGCTATGGGGCTAGA	900
Qy	972	ggaatcaggccaagagttcatttgggtagttaaaaaaggccaaaaacgaagaagaaggaaa	1031
Db	901	AGCCTCTGGACAAGATTTCAATTTGGGTATCA-----GAAC	936
Qy	1032	aggaaaagaagaatggctgccagaaaattttgaggaaagagtgaagatagaggcttgat	1091
Db	937	AGGGAATGAAGATTGGCTCCCAGAAGGATTTCGAGGAAAGAACAAAAGAAAAGGTTTAAT	996
Qy	1092	cataagaggatgggcgccgcaattgttgatactcgatcatcctgcggtaggagctttcgt	1151
Db	997	CATAAGAGGATGGGCACCCCAAGTGTGATTCTTGATCACGAAGCTATTGGAGCTTTTGT	1056
Qy	1152	gacgcattgtggatggaattcgacgcttgaaggaatatgcgcgggtgtgcctatggtgac	1211
Db	1057	TACTCATTGTGGATGGAACCTGACACTGGAAGGAATATCAGCAGGGGTACCAATGGTGAC	1116
Qy	1212	ttggccagtttttcgcagagcagtttttcaatgagaagtttgtgacagaggttttggggac	1271
Db	1117	ATGGCCAGTATTTGCGGAACAGTTTTTCAATGAGAAGTTGGTGACTGAGGTAATGAGAAG	1176
Qy	1272	cggtgtttcggttgggaataagaagtggctaagggcagcaagtgaaggtgtgtcgagggg	1331
Db	1177	TGGAGCTGGTGTGGTTCTAAGCAATGGAAGAGAACAGCTAGTGAAGGAGTGAAAAGAGA	1236
Qy	1332	ggcagtgcgaacgcggtgcagcgtgttatgggtgggagaaaatgcgtcggagatgagaaa	1391
Db	1237	AGCAATAGCAAAGGCGATAAAGAGAGTAATGGCGAGTGAAGAAACAGAGGGATTGAGAAG	1299

Qy	1392	gcgagcgaagtattataaggaaaatggcgaggcggttgaggaaggcggttcgtctta	1451
Db	1297	CAGAGCAAAGAGGTACAAAGAAATGGCAAGAGAAGCTATTGAAGAAGGAGGATCATCTTA	1356
Qy	1452	taatggtttgaatgagatgatagaggatttgagtgtgtaccgtgctccagaaaaacaaga	1511
Db	1357	CAATGGATGGGCTACTTTGATAACAAGACATAACTTCATATCGTTAACTAGTTGATGCAA	1416
Qy	1512	cttaaactagattccttatagatgacttctagtgtagacaattgtaattttttgccttttat	1571
Db	1417	AAAAGAAAAAACATGTGTGTTTCTATATTCTGTCTTCTGTTTTGCTGATTTGATCATATT	1476
Qy	1572	tcaagtttcctcattagtggtga	1594
Db	1477	ACGTACTTCTTCATGATAATTAA	1499

RESULT 1
 US-09-106-464-2
 ; Sequence 2, Application US/09106464
 ; Patent No. 6011145
 ; GENERAL INFORMATION:
 ; APPLICANT: Steffens, John C.
 ; APPLICANT: Ghangas, Gurdev S.
 ; APPLICANT: Kuai, Jian-Ping
 ; APPLICANT: Eannetta, Nancy
 ; TITLE OF INVENTION: Chain Length Specific UDP-Glc:Fatty Acid
 ; TITLE OF INVENTION: Glucosyltransferases
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Jones, Tullar & Cooper, P.C.
 ; STREET: P.O. Box 2266 Eads Station
 ; CITY: Arlington
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/106,464
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/055,554
 ; FILING DATE: 13-AUG-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Spector, Eric S.
 ; REGISTRATION NUMBER: 22495
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-415-1500
 ; TELEFAX: 703-415-1508
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 471 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-106-464-2

Query Match 60.3%; Score 1516.5; DB 3; Length 471;
 Best Local Similarity 58.2%; Pred. No. 5.8e-155;
 Matches 278; Conservative 84; Mismatches 103; Indels 13; Gaps 3;

Qy 1 MGKLHIALFPVMAHGHMIPMLDMAKLFTSRGIQTTIIST---LAFADPINKARDSGLDI 56
 ||:|| ||:|| ||||| ||||| ||:|| ||:|| | : : |:||
 Db 2 MGQLHFFFFPMMAQGHMIPMLDMAKLTVACRGVKATIITPLNESVFSKAIERNKHLGIEI 61
 Qy 57 GLSILKFPPEGSGIPDHMVSLDLV-TEDWLPKFVESLVLLQEPVEKLIEELKLDCLVSDM 115
 : :||| : : :| ||| :| || :|| :|| :|||
 Db 62 DIRLLKFPKENDLPEDCERLDLVPSDDKLPNFLKAAAMMKDEFEEELIGECRPDCLVSDM 121

Qy 116 FLPWTVDCAAKFGIPRLVFHGTSNFALCASEQMKLHKPYKNVTSDTETFVIPDFPHELKF 175
 ||||| | |||| | |||:||||| |||| : :: :||:||||:|||||:|||||:|:
 Db 122 FLPWTTDSAAKFSIPRIVFHGTSYFALCVGDSIRRKNKPFKNVSSDTETFVVPDFPHEIRL 181
 Qy 176 VRTQVAPFQLAETENGFSKLMKQMTESVGRSYGVVNSFYELESTYVDYYREVLGRKSWN 235
 |||::||: :: | | : ::| : || :|||: ||||| ||::| :|:||||:
 Db 182 TRTQLSPFEQSDEETGMAPMIKAVRESDAKSYGVIFNSFYELES DYVEHYTKVVGRKNWA 241
 Qy 236 IGPLLLSNNGNEEKVQRGKESAIGEHECLAWLNSKKQNSVVYVCFGSMATFTPAQLRETA 295
 |||| | | | :||:|:| || || |:|| :|:||||| | | | ||:| |
 Db 242 IGPLSLCNRDIEYKAERGRKSSIDEHACLKWLDSKKSSSIVYVCFGSTADFTTAQMQLA 301
 Qy 296 IGLEESGQEFIWVVKAKNEEEGKGKEEWLPENFEERVKDRGLIIRGWAPQLLILDHPAV 355
 :||| |||:||||:| | | |||| | ||| ||||| ||||| ||||| |:
 Db 302 MGLEASGQDFIWVIR-----TGNE DWLPEGFEERTKEKGLIIRGWAPQVLILDHEAI 353
 Qy 356 GAFVTHCGWNSTLEGICAGVPMVTWVPVFAEQFFNEKFVTEVLGTGVS VGNKKWLRAASEG 415
 ||||| ||||| ||||| |||||:||||| ||||| ||||| :| ||:| | ||||
 Db 354 GAFVTHCGWNSTLEGISAGVPMLTWPVFAEQFFNEKLVTEVMRSGAGVGSKQWKRTASEG 413
 Qy 416 VSREAVTNAVQVRVMGENASEMRKRAKYYKEMARRAVEEGGSSYNGLNEMIEDLSVYR 473
 | |||: |::||| | | || | ||||| :| ||||| ||||| :|:|:| |
 Db 414 VKREAIAKAIKRVMA SEETEGFRSRAKEYKEMAREAIEEGGSSYNGWATLIQDITSYR 471

RESULT 2

US-08-797-226-2

; Sequence 2, Application US/08797226

; Patent No. 5959180

; GENERAL INFORMATION:

; APPLICANT: MOEHS, CHARLES P

; APPLICANT: ALLEN, PAUL V

; APPLICANT: ROCKHOLD, DAVID R

; APPLICANT: STAPLETON, ANDREW

; APPLICANT: GARBARINO, JOAN E

; APPLICANT: FRIEDMAN, MENDEL

; APPLICANT: BELKNAP, WILLIAM R

; TITLE OF INVENTION: DNA SEQUENCES ENCODING SOLANIDINE

; TITLE OF INVENTION: UDP-GLUCOSE GLUCOSYLTRANSFERASE AND USE TO REDUCE

; TITLE OF INVENTION: GLYCOALKALOIDS IN SOLANACEOUS PLANTS

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NANCY J. PARSONS

; STREET: 800 BUCHANAN ST.

; CITY: ALBANY

; STATE: CA

; COUNTRY: USA

; ZIP: 94710

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/797,226

; FILING DATE:

```

; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PARSONS, NANCY J
; REGISTRATION NUMBER: 40,364
; REFERENCE/DOCKET NUMBER: 0011.97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-5731
; TELEFAX: (510) 559-5777
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-226-2

```

```

Query Match          34.9%; Score 877; DB 2; Length 488;
Best Local Similarity 38.1%; Pred. No. 6.7e-86;
Matches 191; Conservative 100; Mismatches 152; Indels 58; Gaps 16;

```

```

Qy      4 LHIALFPVMAHGHMIPMLDMAKLFTSRGIQTTIIST---LAFADPI-NKARDSGLDIGL 58
      ||: | :| |||:: |:| |||:: ||::| | | : | || | :
Db     11 LHVLFPLFSLAGHFIPLVNAARLFASRGVKATILTPHNALLFRSTIDDDVRISGFPISI 70

Qy     59 SILKFPPEGSGIPDHMVSLDLVTEDWLP-KFVESLVLLQEPVEKLIIEELKLDCLVSDMFL 117
      :||| |::: | : | :| | :| |||::|:| | ||: ||: |||:
Db     71 VTIKFPSAEVGLPEGIESFNSATSPEMPHKIFYALSLLQKPMEDKIRELRPDCIFSDMYF 130

Qy    118 PWTVDCAAKFGIPRLVFHGTSNFALCASEQMKLHKPYKNVTS-D-TETFVIPDFPHELKFKV 176
      ||||| | : |||::: : : :|||::| | :|||:| | |::|
Db    131 PWTVDIADDELHIPRILYNLSAYMCYSIMHNLKVYRPHKQPNLDESQSFFVVPGLPDEIKFK 190

Qy    177 RTQVAPFQLAETENG---FSKLMKQMTESVGRSYGVVNSFYELESTYVDYYREVLGRKS 233
      :|: | :||: | :||: | :| |||::| :||| | ||||::: |
Db    191 LSQLTD-DLRKSDDQKTVFDELLEQVEDSEERSYGIVHDTFYELEPAYVDYYQKLKKPKC 249

Qy    234 WNIGPL-----LLSNNNGNEEKVQRGKESAIGEHECLAWLNSKKQNSVVYVCFGS 282
      |: ||| | :| : | | | : |||::| ||:| | ||
Db    250 WHFGPLSHFASKIRSKELISEHNNNEIV-----IDWLNAQKPKSVLYVSFGS 296

Qy    283 MATFTPAQLRETAIGLEESGQEFIWVVKAKNEEEGKGKEEWLP-ENFEERVKDRGLIIR 341
      || | :|| | | | :| |||:: ||| || | :| :|| | :
Db    297 MARFPESQLNEIAQALDASNPFIIFVLR--PNEETA----SWLPVGNLEDKTK-KGLYIK 349

Qy    342 GWAPQLLILDHPAVGAFVTHCGWNSTLEGICAGVPMVTPVFAEQFFNEKFPVTEVLGTGV 401
      || ||| |::| | | :||| | || | | |||::|:|:|:| | | | :
Db    350 GWVPQLTIMEHSATGGFMTHCGTNSVLEAITFGVPMITWPLYADQFYNEK-VVEVRGLGI 408

Qy    402 SVGNKKWLRAASEG-----VSREAVTNAVQRMVGENASE---MRKRAKYYKEMARRA 451
      :| | :|| : : : :||: | :| :| :||: |
Db    409 KIGIDVW---NEGIEITGPVIESAKIREAIERLMISNGSEEIINIRDRVMAMSKMAQNA 464

Qy    452 VEEGGSSYNGLNEMIEDLSVY 472
      |||||:| | :|: : |
Db    465 TNEGGSSWNNLTALIQHIKNY 485

```


Qy	134	tgttggacatgggccaagctcttttacctcaagaggcatatacaaaacaacaatcatttcgactc	193
Db	62	CACTTGACATGGCGAAGCTTGTCGCTTGTCGTGGTGTTAAAGCCACTATAATCACAACAC	121
Qy	194	tcgcc-----ttcgcgtgatccgataaacaagctcgtgattcgggcctcgata	241
Db	122	CTCTCAATGAATCTGTTTTCTCTAAAGCTATTGAGAGAAACAAGCATTTAGGTATTGAAA	181
Qy	242	ttggactaagcatcctcaaattcccaccagaaggatcaggaataccagatcacatggtga	301
Db	182	TTGATATTTCGTTTACTAAAAATCCCAGCTAAGGAGAATGATTGTCCTGAAGATTGTGAGC	241
Qy	302	gccttgatctagt---tactgaagattggctcccaaagtttggtgagtcattagtcttat	358
Db	242	GTCTTGATCTTGTACCTTCTGATGACAAACTCCCAAACCTCTTAAAGCTGCGGCTATGA	301
Qy	359	tacaagagccagttgagaagcttatcgaagaactaaagctcgactgtctcgtttccgaca	418
Db	302	TGAAAGATGAATTTGAGGAGCTTATTGGAGAATGTCGCCCTGATTGTCTTGTTTCTGATA	361
Qy	419	tgttcttgcttggacagtcgattgtgcggttaagttcggatttcgaggttggttttcc	478
Db	362	TGTTCCCTCCATGGACTACTGATAGTGCAGCCAAATTTAGCATACCAAGAATTGTATTCC	421
Qy	479	acggaacgagcaactttgcgttggtgtgcttcggagcaaataagcttcacaagccttata	538
Db	422	ATGGAAGTAGTTACTTTGCGCTTTGTGTTGGCGATAGCATCAGGCGTAATAAGCCTTTCA	481
Qy	539	agaatgtaacttctgatactgagacatttggtataaccggatttcccgcagtgagctgaagt	598
Db	482	AGAATGTGTCATCGGATACTGAAACTTTTGTGTACCGGATTTTCCACATGAAATTAGGC	541
Qy	599	ttgtgaggactcaagtggctccgtttcagcttgcggaacggagaatggattctcaaagt	658
Db	542	TAACTAGAACACAGTTGTCTCCGTTTGAGCAATCGGATGAAGAGACGGGTATGGCTCCCA	601
Qy	659	tgatgaaacagatgacggagtcgttggtagaagctacggtggttggttaacagttttt	718
Db	602	TGATTAAAGCTGTGAGGGAATCGGATGCGAAGAGCTATGGAGTTATATTCAATAGCTTTT	661
Qy	719	atgagctcgagtcgacttatgtggattattacagagaggttttggttagaaagtcttgga	778
Db	662	ATGAGCTTGAATCAGATTATGTTGAACATTACACTAAGGTTGTAGGTAGAAAAAATTGGG	721
Qy	779	atatagggcctctgttggttatccaacaatggcaatgaggaaaaagtacaaaggggaaagg	838
Db	722	CTATTGGTCCGCTTTCGCTGTGCAATAGGGATATTGAATATAAAGCGGAAAGAGGGAGGA	781
Qy	839	aatctgcgattggcgaacacgaatgcttggttggttgattccaagaagcagaattcgg	898
Db	782	AATCATCTATCGATGAACACGCGTGCTTGAAATGGCTTGATTCTGAAGAAATCAAGTTCCA	841
Qy	899	ttgtttacgtttgttttggaagtatggcgacttttactccagcgcagttgcgcgaaactg	958
Db	842	TTGTTTATGTTTGTGTTTGGGAAGTACAGCAGATTTCACTACAGCACAGATGCAAGAACTTG	901

Qy 959 cgattggactcgaggaatcaggccaagagttcatttgggtagttaaaaaggccaaaaacg 1018
 | | | | | | | | | | | | | | | | | | | | | |
 Db 902 CTATGGGGCTAGAAGCCTCTGGACAAGATTTTCATTGGGTATCA----- 946

 Qy 1019 aagaagaaggaaaaggaaaagaagaatggctgccagaaaattttgaggaaagagtgaaag 1078
 | | | | | | | | | | | | | | | | | | | | | |
 Db 947 -----GAACAGGGAATGAAGATTGGCTCCCAGAAGGATTCGAGGAAAGAACAAAAG 997

 Qy 1079 atagaggcttgatcataagaggatggcgccgcaattgttgatactcgatcatcctgcgg 1138
 | | | | | | | | | | | | | | | | | | | | | |
 Db 998 AAAAAGGTTTAATCATAAGAGGATGGGCACCCCAAGTGCTGATTCTTGATCACGAAGCTA 1057

 Qy 1139 taggagcttttcgtgacgcattgtggatggaattcgacgttggaaggaatatgcgccggtg 1198
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1058 TTGGAGCTTTTGTACTCATTGTGGATGGAACCTCGACACTGGAAGGAATATCAGCAGGGG 1117

 Qy 1199 tgcctatggtgacttggccagttttcgcagagcagtttttcaatgagaagtttgtgacag 1258
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1118 TACCAATGTTGACATGGCCAGTATTTGCGGAACAGTTTTTCAATGAGAAGTTGGTGAAGT 1177

 Qy 1259 aggttttggggaccggtgtttcggttgggaataagaagtggctaagggcagcaagtgaag 1318
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1178 AGGTAATGAGAAGTGGAGCTGGTGTGGTTCTAAGCAATGGAAGAGAACAGCTAGTGAAG 1237

 Qy 1319 gtgtgtcgaggggagggcagtgacgaacgcggtgcagcgtgttatggtgggagaaaatgcgt 1378
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1238 GAGTGAAAAGAGAAGCAATAGCAAAGGCGATAAAGAGAGTAATGGCGAGTGAAGAAACAG 1297

 Qy 1379 cggagatgagaaagcgagcggaagtattataaggaaatggcgaggcggcggttgaggaag 1438
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1298 AGGGATTGAGAAGCAGAGCAAAAGAGTACAAAGAAATGGCAAGAGAAGCTATTGAAGAAG 1357

 Qy 1439 gcggttcgtcttataatggtttgaatgagatgatagaggatttgagtgtgtaccgtgctc 1498
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1358 GAGGATCATCTTACAATGGATGGGCTACTTTGATACAAGACATAACTTCATATCGTTAAC 1417

 Qy 1499 cagaaaaacaagacttaaaactagattcttatagatgacttctagtgtgacaattgtaatt 1558
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1418 TAGTGATGCAAAAAAAGAAAAAACATGTGTGTTTCTATATTCTGTCTTCTGTTTTGCTG 1477

 Qy 1559 ttttgcctttttattcaagttttcctcattagtggtga 1594
 | | | | | | | | | | | | | | | | | |
 Db 1478 ATTTGATCATATTACGTACTTCTTCATCATAATTAA 1513

L16 ANSWER 1 OF 1 AGRICOLA

ACCESSION NUMBER: 97:35254 AGRICOLA
DOCUMENT NUMBER: IND20566206
TITLE: Identification of an immediate-early salicylic acid-inducible tobacco gene and characterization of induction by other compounds.

AUTHOR(S): Horvath, D.M.; Chua, N.H.
CORPORATE SOURCE: The Rockefeller University, New York, NY.
AVAILABILITY: DNAL (QK710.P62)
SOURCE: **Plant molecular biology, Aug 1996.**

Vol. 31, No. 5. p. 1061-1072
Publisher: Dordrecht : Kluwer Academic Publishers.
CODEN: PMBIDB; ISSN: 0167-4412

NOTE: Includes references

PUB. COUNTRY: Netherlands

DOCUMENT TYPE: Article

FILE SEGMENT: Non-U.S. Imprint other than FAO

LANGUAGE: English

AB Tobacco genes that are induced in response to salicylic acid (SA) treatment with immediate-early kinetics were identified by differential mRNA display. Detailed analysis of IS10a, one cDNA clone identified by this method, revealed induction within 30 min of treatment, with a peak

of expression at 3 h, that decayed rapidly thereafter. Treatment with the protein synthesis inhibitor, cycloheximide (CHX), also caused induction

of IS10a mRNA to comparable levels, but the IS10a mRNA continued to accumulate after 3 h of induction. In combination, CHX and SA led to a superinduction of IS10a mRNA levels that was also sustained. Half-maximal induction was evident at ca. 100-150 micromolar SA. In addition to SA, induction of IS10a occurred to varying degrees upon treatment with acetylsalicylic acid, benzoic acid, 2,4-dichlorophenoxyacetic acid,

methyl jasmonate, and hydrogen peroxide, whereas treatment with other compounds had no effect. The proteins encoded by IS10a and a second highly homologous cDNA show sequence similarity to UDP-glucose: flavonoid glucosyltransferases.

RESULT 2
 NTU32644
 LOCUS NTU32644 1624 bp mRNA PLN 25-NOV-1996
 DEFINITION *Nicotiana tabacum* immediate-early salicylate-induced glucosyltransferase (IS5a) mRNA, complete cds.
 ACCESSION U32644
 VERSION U32644.1 GI:1685004
 KEYWORDS .
 SOURCE common tobacco.
 ORGANISM *Nicotiana tabacum*
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; *Nicotiana*.
 REFERENCE 1 (bases 1 to 1624)
 AUTHORS Horvath,D.M. and Chua,N.H.
 TITLE Identification of an immediate-early salicylic acid-inducible tobacco gene and characterization of induction by other compounds
 JOURNAL Plant Mol. Biol. 31 (5), 1061-1072 (1996)
 MEDLINE 97000918
 REFERENCE 2 (bases 1 to 1624)
 AUTHORS Horvath,D.M. and Chua,N.H.
 TITLE Direct Submission
 JOURNAL Submitted (29-JUL-1995) Diana M. Horvath, Laboratory of Plant Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA
 FEATURES Location/Qualifiers
 source 1. .1624
 /organism="Nicotiana tabacum"
 /strain="bright yellow 2"
 /db_xref="taxon:4097"
 gene 101. .1531
 /gene="IS5a"
 CDS 101. .1531
 /gene="IS5a"
 /codon_start=1
 /product="immediate-early salicylate-induced glucosyltransferase"
 /protein_id="AAB36653.1"
 /db_xref="GI:1685005"
 /translation="MGQLHIFFFPVMAGHMIPTLDMAKLFASRGVKATIITPLNEF
 VFSKAIQRNKHGLGIEIEIRLIKFPVAVENGLPEECERLDQIPSEKLPNFFKAVAMMQE
 PLEQLIEECRPDCLISDMFLPWTTDTAAKFNIPIRVFHGTSFFALCVENSURLNKPFFK
 NVSSDSETFVVPDLPHEIKLRTQVSPFERSGEETAMTRMIKTVRESDSKSYGVVFNS
 FYELETDYVEHYTKVLGRRWAIGPLSMCNRDIEDKAERGKSSIDKHECLKWLDSSK
 PSSVVYICFGSVANFTASQLHELAMGVEASGQEFIWVVRTELDNEDWLPEGFEERTKE
 KGLIIRGWAPQVLILDHESVGAFVTHCGWNSTLEGVSGGVPMVTWPVFAEQFFNEKLV
 TEVLKTGAGVGSIQWKRSASEGVKREAIKAIKRMVSEEDGFRNRAKAYKEMARKA
 IEEGGSSYTGLTTLLLEDISTYSSTGH"
 BASE COUNT 498 a 293 c 365 g 468 t
 ORIGIN

Query Match 33.6%; Score 588.2; DB 14; Length 1624;
 Best Local Similarity 64.4%; Pred. No. 1.3e-128;
 Matches 944; Conservative 0; Mismatches 488; Indels 33; Gaps 3;

Qy	68	ttttaaaaaatgggaaaacttcacattgccttattttccagttatggctcatggtcacatga	127
Db	93	TTTCACAAATGGGTCAGCTCCATATTTTCTTCTTTCTGTGATGGCTCATGGCCACATGA	152
Qy	128	tcccaatggttgacatggccaagctctttacctcaagaggcatacaaaacaacaatcattt	187
Db	153	TTCCTACACTAGACATGGCGAAGCTCTTTGCTTCACGTGGTGTTAAGGCCACTATAATCA	212
Qy	188	cgactc-----tcgccttcgctgatccgataaacaagctcgtgattcggggc	235
Db	213	CAACCCCACTCAATGAATTCGTTTTCTCCAAAGCTATTCAAAGAAACAAGCATTTGGGTA	272
Qy	236	tcgatattggactaagcatcctcaaattcccaccagaaggatcaggaataccagatcaca	295
Db	273	TCGAAATCGAAATCCGTTTGATCAAATTCCCAGCTGTTGAAAACGGCTTACCTGAAGAAT	332
Qy	296	tggtgagccttgatctagt---tactgaagattggctcccaaagtttggtgagtcattag	352
Db	333	GCGAACGCCTCGATCAAATCCCTTCAGATGAGAAGCTCCCAAACCTTTTCAAAGCTGTAG	392
Qy	353	tcttattacaagagccagttgagaagcttatcgaagaactaaagctcgactgtctcgttt	412
Db	393	CTATGATGCAAGAACCACTAGAACAGCTTATTGAAGAATGTGCGCCCGATTGTCTTATTT	452
Qy	413	ccgacatggttcttgaccttgacagtcgattgtgcggttaagttcggtattccgaggttg	472
Db	453	CAGATATGTTTCCTTCCTTGGACTACTGATACTGCAGCAAAATTTAACATTCCAAGAATAG	512
Qy	473	ttttccacggaacgagcaactttgcggttggtgcttcggagcaaataagcttcacaagc	532
Db	513	TCTTTTCATGGCACAAGCTTCTTTGCTCTTTGTGTTGAGAATAGCGTCAGGCTAAATAAGC	572
Qy	533	cttataagaatgtaacttctgatactgagacatttggtataaccggatttcccgcgatgac	592
Db	573	CTTTCAAGAATGTGTCTCAGATTCTGAAACTTTTGTGTACCGGATTTGCCTCACGAAA	632
Qy	593	tgaagtttgtaggactcaagtggctccgtttcagcttgcggaacggagaatggattct	652
Db	633	TTAAGCTGACCAGAACCCAGGTGTCTCCGTTTGAGCGATCTGGGGAAGAGACGGCTATGA	692
Qy	653	caaagttgatgaaacagatgacggagtctgttggtagaagctacggtggttggttaaca	712
Db	693	CCCGGATGATAAAACAGTCAGGGAATCAGATTCAAAGAGCTATGGAGTTGTTTTCAACA	752
Qy	713	gttttttatgagctcgagtcgacttatgtggattattacagagaggttttgggtagaaagt	772
Db	753	GTTTCTATGAGCTTGAAACAGATTATGTTGAGCATTATACTAAGGTGCTGGGTAGAAGAG	812
Qy	773	cttggaatatagggcctctgttggttatccaacaatggcaatgaggaaaaagtacaaaggg	832
Db	813	CTTGGGCTATTGGCCCTCTATCGATGTGCAACAGGGACATTGAAGATAAAGCTGAAAGAG	872
Qy	833	gaaaggaatctgcgattggcgaacacgaatgcttggcttggttgaattccaagaagcaga	892
Db	873	GAAAGAAATCCTCTATTGATAAACACGAGTGCTTGAAATGGCTTGATTCTGAAGAAACCAA	932
Qy	893	attcgqttgtttacgqtttgttttggaaagtatggcgacttttactccagcgagttgcgcg	952

Db	933	GTTCCGTCGTTTACATTTGTTTTGGAAGCGTAGCGAATTTCACTGCATCACAACCTGCACG	992
Qy	953	aaactgcgattggactcgaggaatcaggccaagagttcatttgggtagttaaaaaggcca	1012
Db	993	AACTTGCTATGGGAGTTGAAGCTTCCGGACAAGAATTCATTTGGGTGTGTAGAA-----	1046
Qy	1013	aaaacgaagaagaaggaaaaggaaaagaagaatggctgccagaaaattttgaggaaagag	1072
Db	1047	-----CAGAACTAGACAACGAAGATTGGTTGCCTGAAGGATTCGAGGAAAGAA	1094
Qy	1073	tgaagatagaggcttgatcataagaggatggcgccgcaattgttgatactcgatcatc	1132
Db	1095	CGAAAGAGAAAGGTTTAATAATAAGAGGATGGGCACCCAAGTACTAATTCTTGATCACG	1154
Qy	1133	ctgcggtaggagcttttcgtgacgcattgtggatggaattcgacgttgggaaggaatatgcg	1192
Db	1155	AATCTGTGGGAGCTTTTGTTACACATTGTGGTTGGAATTCAACACTAGAAGGAGTTTCAG	1214
Qy	1193	ccggtgtgcctatggtgacttggccagttttcgcagagcagttttcaatgagaagtttg	1252
Db	1215	GAGGGGTTCCAATGGTAACATGGCCTGTATTTGCTGAGCAATTTTTCATGAGAAGTTAG	1274
Qy	1253	tgacagaggttttggggaccggtgtttcggttgggaataagaagtggctaagggcagcaa	1312
Db	1275	TGACTGAGGTTTTGAAAACCTGGAGCTGGTGTGGTTTCGATACAATGGAAGAGATCAGCTA	1334
Qy	1313	gtgaaggtgtgtcagaggaggcagtgacgaacgcggtgcagcgtgttatggtgggagaaa	1372
Db	1335	GTGAAGGAGTGAAAAGAGAAGCAATAGCTAAGGCAATAAAGAGAGTAATGGTGAGTGAAG	1394
Qy	1373	atgcgtcggagatgagaaagcgagcgaagtattataaggaaatggcgaggcgggcggttg	1432
Db	1395	AAGCAGATGGATTTCAGAAACAGAGCTAAAGCGTATAAGGAGATGGCAAGAAAGGCTATTG	1454
Qy	1433	aggaaggcgggttcgtcttataatggtttgaatgagatgatagaggatttgagtgtgtacc	1492
Db	1455	AAGAAGGAGGGTCATCTTACACTGGATTGACTACTTTGTTGGAAGATATAAGTACATATA	1514
Qy	1493	gtgctccagaaaaacaagacttaaa	1517
Db	1515	GTTCCACTGGTCATTAAGTTATGAA	1539

RESULT 5
 AF346431
 LOCUS AF346431 1428 bp mRNA PLN 02-APR-2001
 DEFINITION *Nicotiana tabacum* phenylpropanoid:glucosyltransferase 1 (togt1) mRNA, partial cds.
 ACCESSION AF346431
 VERSION AF346431.1 GI:13492673
 KEYWORDS .
 SOURCE common tobacco.
 ORGANISM *Nicotiana tabacum*
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; *Nicotiana*.
 REFERENCE 1 (bases 1 to 1428)
 AUTHORS Fraissinet-Tachet, L., Baltz, R., Chong, J., Kauffmann, S., Fritig, B. and Saindrenan, P.
 TITLE Two tobacco genes induced by infection, elicitor and salicylic acid encode glucosyltransferases acting on phenylpropanoids and benzoic acid derivatives, including salicylic acid
 JOURNAL ~~PLoS Pathog. 4(3): 319-323 (1998)~~
 MEDLINE 99039922
 PUBMED 9824316
 REFERENCE 2 (bases 1 to 1428)
 AUTHORS Fraissinet-Tachet, L., Baltz, R., Chong, J., Fritig, B., Beffa, R. and Saindrenan, P.
 TITLE Direct Submission
 JOURNAL Submitted (05-FEB-2001) Phytopathologie Moleculaire, Institut de Biologie Moleculaire des Plantes du CNRS, 12 rue du general Zimmer, Strasbourg 67000, France
 FEATURES
 Location/Qualifiers
 source 1. .1428
 /organism="Nicotiana tabacum"
 /cultivar="Samsun NN"
 /db_xref="taxon:4097"
 gene 1. .>1428
 /gene="togt1"
 /note="allele of *Nicotiana tabacum* IS5a encoded by GenBank Accession Number U23643"
 CDS 1. .>1428
 /gene="togt1"
 /note="glucosyltransferase"
 /codon_start=1
 /product="phenylpropanoid:glucosyltransferase 1"
 /protein_id="AAK28303.1"
 /db_xref="GI:13492674"
 /translation="MGQLHFFFPVMAHGHIPTLDMAKLFASRGVKATIITTPLENEF
 VFSKAIQRNKHLLGIEIEIRLIKFPVAVENGLPEECERLDQIPSDKLPNFFKAVAMMQE
 PLEQLIEECRPDCLISDMFLPWTTDTAAKFNIPIRVFHGTSFFALCVENSVRNKPFFK
 NVSSDSETFVVPDLPHEIKLRTQVSPFERSGEETAMTRMIKTVRESDSKSYGVVFNS
 FYELETDYVEHYTKVLGRRAWAIGPLSMCNRDIEDKAERGKSSIDKHECLKWLDSSK
 PSSVYVCFGSVANFTASQLHELAMGIEASQEFIWVVRTELDNEDWLPEGFEERTKE
 KGLIIRGWAPQVLILDHESVGAFTVTHCGWNSTLEGVSGGVPMVTWPVFAEQFFNEKLV
 TEVLKTGAGVGSIQWKRSASEGKREAIKAIKRVVMVSEADGFRNRKAYKEMARKA
 IEEGSSYTGLTTLLEDISTYSSTGH"
 BASE COUNT 440 a 256 c 340 g 392 t
 ORIGIN

Query Match 33.2%; Score 581.8; DB 13; Length 1428;
Best Local Similarity 64.6%; Pred. No. 4.1e-127;
Matches 931; Conservative 0; Mismatches 477; Indels 33; Gaps 3;

```
Qy 76 atgggaaaacttcacattgccttatttccagttatggctcatggtcacatgatcccaatg 135
    ||||| | || | | || | || | || | || | || | || | || | || | || |
Db 1 ATGGGTCAGCTCCATTTTTTCTTCTTTCTGTGATGGCTCATGGCCACATGATTCCTACA 60

Qy 136 ttggacatggccaagctctttacctaagaggcatacaaaacaacaatcatttcgactc-- 193
    | ||||| | ||||| | || | | || | | || | | || | | || | | || |
Db 61 CTAGACATGGCGAAGCTCTTTGCTTCACGTGGTGTTAAGGCCACTATAATCACAACCCCA 120

Qy 194 -----tcgccttcgctgatccgataaacaagctcgtgattcgggcctcgatatt 243
    || | || | || | || | || | || | || | || | || | || | || |
Db 121 CTCAATGAATTGTTTTCTCCAAAGCTATTCAAAGAAACAAGCATTGGGTATCGAAATC 180

Qy 244 ggactaagcatcctcaaattcccaccagaaggatcaggaataccagatcacatgggtgagc 303
    | | | | | ||||| ||||| | | | | | || | || | || | || | || |
Db 181 GAAATCCGTTTGATCAAATTTCCAGCTGTTGAAACGGCTTACCTGAAGAATGCGAACGC 240

Qy 304 cttgatctagt---tactgaagattggctcccaaagtttggtgagtcattagctcttatta 360
    || |||| | | | | | || | || | || | || | || | || | || | || |
Db 241 CTCGATCAAATCCCTTCAGATGAGAAGCTCCCAAACCTTTTTCAAAGCTGTAGCTATGATG 300

Qy 361 caagagccagttgagaagcttatcgaagaactaaagctcgactgtctcgtttccgacatg 420
    ||||| || | || | ||||| ||||| | || | || | || | || | || | || |
Db 301 CAAGAACCACTAGAACAGCTTATTGAAGAATGTCGCCCGATTGTCTTATTTAGATATG 360

Qy 421 ttcttgcccttgacagtcgattgtgcggctaagttcggtattccgaggttggtttccac 480
    || | ||||| || | || | || | || | || | || | || | || | || | || |
Db 361 TTCCTTCCTTGACTACTGATACTGCAGCAAAATTTAACATTCCAAGAATAGTCTTTCAT 420

Qy 481 ggaacgagcaactttgcgttggtgcttcggagcaaatgaagcttcacaagccttataag 540
    || | || | ||||| ||||| | | | | | || | || | || | || | || | || |
Db 421 GGCACAAGCTTCTTTGCTCTTTGTGTTGAGAATAGCGTCAGGCTAAATAAGCCTTTCAAG 480

Qy 541 aatgtaacttctgatactgagacatttggtataccggatttcccgcatgagctgaagttt 600
    ||||| | | | || | || | || | || | || | || | || | || | || | || |
Db 481 AATGTGTCCTCAGATTCTGAAACTTTTGTTGTACCGGATTTGCCTCACGAAATTAAGCTG 540

Qy 601 gtgaggactcaagtggctccgtttcagcttgcggaacggagaatggattctcaaagttg 660
    || | || | || | ||||| ||||| | | | || | || | || | || | || | || |
Db 541 ACCAGAACCCAGGTGTCTCCGTTTGAGCGATCTGGGGAAGAGACGGCTATGACCCGGATG 600

Qy 661 atgaaacagatgacggagtcgttggttagaagctacggtggttggttaacagtttttat 720
    || || | | | || | || | || | || | || | || | || | || | || | || |
Db 601 ATAAAAACAGTCAGGGAATCAGATTCAAAGAGCTATGGAGTTGTTTTCAACAGTTTCTAT 660

Qy 721 gagctcgagtcgacttatgtggattattacagagaggttttgggtagaaagctcttgaat 780
    ||||| || | ||||| || | || | || | || | || | || | || | || | || |
Db 661 GAGCTTGAAACAGATTATGTTGAGCATTATACTAAGGTGCTGGGTAGAAGAGCTTGGGCT 720

Qy 781 atagggcctctgttggtatccaacaatggcaatgaggaaaaagtacaaaggggaaaggaa 840
    || | || ||||| | | | ||||| | | | || | || | || | || | || | || |
Db 721 ATTGGCCCTCTATCGATGTGCAACAGGGACATTGAAGATAAAGCTGAAAGAGGAAAGAAA 780
```

Qy 841 tctgcgattggcgaacacgaatgcttggcttgggtgaattccaagaagcagaattcggtt 900
 || | ||| | ||||| | |||| | || | |||| | |||| | || ||||
 Db 781 TCCTCTATTGATAAACACGAGTGCTTGAAATGGCTTGATTCTGAAGAAACCAAGTTCCGTC 840

Qy 901 gtttacgtttgttttgaagtatggcgacttttactccagcgcagttgcgcgaaactgcg 960
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 841 GTTTACGTTTGTGTTTGAAGCGTAGCGAATTTCACTGCATCACAACTGCACGAACCTTGCT 900

Qy 961 attggactcgaggaatcaggccaagagttcatttgggtagttaaaaaggccaaaaacgaa 1020
 || | || | || | || | || | || | || | || | || | || | || | || | || |
 Db 901 ATGGGAATTGAAGCTTCCGGACAAGAATTCATTTGGGTTGTTAGAA----- 946

Qy 1021 gaagaaggaaaaggaaaagaagaatggctgccagaaaattttgaggaaagagtgaagat 1080
 || | || | || | || | || | || | || | || | || | || | || | || | || |
 Db 947 ----CAGAACTAGACAACGAAGATTGGTTGCCTGAAGGATTCGAGGAAAGAACGAAAGAG 1002

Qy 1081 agaggcttgatcataagaggatggcgccgcaattgttgatactcgatcatcctgcggtta 1140
 | || | || | || | || | || | || | || | || | || | || | || | || | || |
 Db 1003 AAAGGTTTAATAATAAGAGGATGGGCACCCCAAGTACTAATTCTTGATCACGAATCTGTG 1062

Qy 1141 ggagctttcgtgacgcattgtggatggaattcgacgttggaaggaatatgcccgggtgtg 1200
 ||||| || | || | || | || | || | || | || | || | || | || | || | || |
 Db 1063 GGAGCTTTTGTTACACATTGTGGTTGGAATTCAACACTAGAAGGAGTTTCAGGAGGGGTT 1122

Qy 1201 cctatgggtgacttggccagttttcgcagagcagtttttcaatgagaagtttgtgacagag 1260
 || | || | || | || | || | || | || | || | || | || | || | || | || |
 Db 1123 CCAATGGTAACATGGCCTGTATTTGCTGAGCAATTTTTCAATGAGAAGTTAGTGACTGAG 1182

Qy 1261 gttttggggaccggtgtttcggttgggaataagaagtggttaagggcagcaagtgaaggt 1320
 ||||| || | || | || | || | || | || | || | || | || | || | || | || |
 Db 1183 GTTTTGAAAACCTGGAGCTGGTGTGTTGCTGATACAATGGAAGAGATCAGCTAGTGAAGGA 1242

Qy 1321 gtgtcgagggaggcagtgacgaacgcggtgcagcgtgttatggtgggagaaaatgcgtcg 1380
 || | || | || | || | || | || | || | || | || | || | || | || | || |
 Db 1243 GTGAAAAGAGAAGCAATAGCTAAGGCAATAAAGAGAGTAATGGTGAGTGAAGAAGCAGAT 1302

Qy 1381 gagatgagaaagcgagcgaagtattataaggaaatggcgaggcggcggttgaggaaggc 1440
 | || | || | || | || | || | || | || | || | || | || | || | || | || |
 Db 1303 GGATTTCAGAAACAGAGCTAAAGCGTATAAGGAGATGGCAAGAAAGGCTATTGAAGAAGGA 1362

Qy 1441 ggttcgtcttataatggtttgaatgagatgatagaggatttgagtgtgtaccgtgctcca 1500
 || | || | || | || | || | || | || | || | || | || | || | || | || |
 Db 1363 GGGTCATCTTACACTGGATTGACTACTTTGTTGGAAGATATAAGTACATATAGTTCCACT 1422

Qy 1501 g 1501
 |
 Db 1423 G 1423

RESULT 1

T03747

glucosyltransferase IS5a (EC 2.4.1.-), salicylate-induced - common tobacco

C:Species: *Nicotiana tabacum* (common tobacco)

C;Date: 24-Mar-1999 #sequence revision 24-Mar-1999 #text change 21-Jul-2000

C;Accession: T03747

R; Horvath, D.M.; Chua, N.H.

Plant Mol. Biol. 31, 1061-1072, 1996

A;Title: Identification of an immediate-early salicylic acid-inducible tobacco gene and characterization of induction by other compounds.

A:Reference number: Z15050; MUID:97000918

A:Accession: T03747

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-476 <HOR>

A;Cross-references: EMBL:U32644; NID:g1685004; PIDN:AAB36653.1; PID:g1685005

A; Experimental source: strain Bright Yellow 2

C: Genetics:

A; Gene: IS5a

C:Superfamily: flavonol O3-glucosyltransferase

C; Keywords: glycosyltransferase; hexosyltransferase

Query Match 61.9%; Score 1558.5; DB 2; Length 476;
Best Local Similarity 59.7%; Pred. No. 2.6e-112;
Matches 285; Conservative 84; Mismatches 97; Indels 11; Gaps 3;

QY 1 MGKHLHIALFPVMAHGHMIPMLDMAKLFSTRGIQTTIIST---LAFADPINKARDSGLDI 56
||:||||| ||||| ||||| |||:| |||:| | : | : : ||:|

Db 1 MGOLHIFFFPVMAHGHMIPTLDMAKLFASRGVKATIITTPLNEFVFSKAIQRNKHGLGIEI 60

Qy 57 GLSILKFPPEGSGIPDHMVSLDLV-TEDWLPKFVESLVLLQEPVEKLIIEELKLDCLVSDM 115
: ::||| :||: || : :: || | :: :|||:|:|||| : |||:|
Db 61 EIRLIKFPVAVENGLPEECERLDAQIPSDEKLPNFFKAVAMMQEPLEQLIEECRPDCLISDM 120

QY 116 FLPWTVDCAAKFGIPRLVFHGTSNFALCASEQMKLHKPYKNVTSDETETFPVIPDFPHELKF 175
| | | | | | | | | | : | | | | | | | | | | : | : | | : | | : | | : | | | : |
Db 121 FLPTTTDTAAKFNIPRIVFHGTSFFALCVENSVRNLNPKPFKNVSSDSETFVVPDLPEIKL 180

Qy 176 VRTQVAPFQLAETENGFSKLMQMTESVGRSYGVVNSFYELESTYVDYYREVLGRKSWN 235
 ||||:|: : | :::| : || :|||| |||||: ||:| :|||:|
 Db 181 TRTOVSPFERSGEETAMTRMIKTVRESDSKSYGVVNSFYELETDYVEHYTKVLGRRAWA 240

QY 236 IGPLLLSNNNGNEEKVQRGKESAIGEHECLAWLNSKKQNSVVYVCFGSMATFTPAQLRETA 295
 |||| : | :| :||:| : |||| ||:| | :|||:||||:| ||:| | |
 Db 241 IGPLSMCNRDIEDKAERGKKSSIDKHECLKWLDSKKPSSVVYICFGSVANFTASOLHELA 300

QY 296 IGLEESGQFEIWWVKKAKNEEEGKGKEEWLPENFEERVKDRGLIIRGWAPQLLILDHPAV 355
:|:| | | | | | | | : : | : | | | | | | | : | | | | : |
Db 301 MGVEASGOEFIWVVRTELD-----NEDWLPEGFEERTKEKGLIIRGWAPOVLILDHESV 354

QY 356 GAFVTHCGWNSTLEGICAGVPMVTWPVFABEQFFNEKFVTEVLGTGVSVGNNKKWLRASEG 415
| | | | | : | | | | | | | | | | | | | | | | : : | : | |
Db 355 GAFVTHCGWNSTLEGVSGGVPMTWPVFABEQFFNEKLVTGLKTGAGVGSIQWKRSASEG 414

QY 416 VSREAVTNAVQVRVMVGENASEMRKRAKYYKEMARRAVEEGSSYNGLNEMIEDLSVY 472

Db 415 VKREAIKAIKRVMSVSEADGFRNRAKAYKEMARKAIEEGSSYTGLTTLLEDISTY 471

RESULT 2

T03745

glucosyltransferase IS10a (EC 2.4.1.-), salicylate-induced - common tobacco

C;Species: Nicotiana tabacum (common tobacco)

C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000

C;Accession: T03745

R;Horvath, D.M.; Chua, N.H.

Plant Mol. Biol. 31, 1061-1072, 1996

A;Title: Identification of an immediate-early salicylic acid-inducible tobacco gene and characterization of induction by other compounds.

A;Reference number: Z15050; MUID:97000918

A;Accession: T03745

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-476 <HOR>

A;Cross-references: EMBL:U32643; NID:g1685002; PIDN:AAB36652.1; PID:g1685003

A;Experimental source: strain Bright Yellow 2

C;Genetics:

A;Gene: IS10a

C;Superfamily: flavonol O3-glucosyltransferase

C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 60.3%; Score 1517.5; DB 2; Length 476;
Best Local Similarity 58.5%; Pred. No. 3.7e-109;
Matches 279; Conservative 85; Mismatches 102; Indels 11; Gaps 3;

```
Qy      1 MGKLHIALFPVMAHGHIPLMDMAKLFTSRGIQTTIIST---LAFADPINKARDSGLDI 56
      ||:||| ||||| ||||| ||||| |||:: |||:| | : | : | ::|
Db      1 MGQLHFFFFFPVMAHGHIPTLMDAKLVASRGVKATIITPLNESVFSKSIQRNKLGLIEI 60

Qy     57 GLSILKFPPEGSGIPDHMVSLDLV-TEDWLPKFVESLVLLQEPVEKLIEELKLDCLVSDM 115
      : ::||| :|:| : |||: ::| || | :: : |||:|:| ||| : :||| ||
Db     61 EIRLIKFPVAVENGLPEECERLDLIPSDDKLPNFFKAVAMMQEPLEQLIEECRPNCVSDM 120

Qy    116 FLPWTVDCAAKFGIPRLVFHGTSNFALCASEQMKLHKPYKNVTSDETTFVIPDFPHELFK 175
      ||||| | |||| :|:| ||||| |||| : ::|:|:|:|:|:|:|:|:|:|:|:|:|
Db    121 FLPWTTDTAAKFNMPRIVFHGTSFFALCVENSIRLNKPFKNVSSDSETFVVPNLPHEIKL 180

Qy    176 VRTQVAPFQLAETENGFSKLMKQMTESVGRSYGVVVNSFYELESTYVDYYREVLGRKSWN 235
      |||:|:|:| : | : ::| : || : |||:| ||| ||| ||:| : |||:|:|
Db    181 TRTQLSPFEQSGETTMTRMIKSVRESDSKSYGVIFNSFNELEHDIYVEHYTKVLGRRAWA 240

Qy    236 IGPLLLSNNGNEEKVQRGKESAIGEHECLAWLNSKKQNSVVYVCFGSMATFTPAQLRETA 295
      |||| : | :| :|:|:|:| :||| ||:| | :||| |||||:| || :| | |
Db    241 IGPLSMCNDRDIEDKAERGKQSSIDKHECLKWLDSKKPSSVVYVCFGSVANFTASQLHELA 300

Qy    296 IGLEESGQEFIWVVKKAKNEEEGKGKEEWLPENFEERVKDRGLIIRGWAPQLLILDHPAV 355
      :|:| ||||| |||: : : |||| ||| |:| ||||| |||||:| ||| :|
Db    301 MGIEASGQEFIWVVRTELD-----NEDWLPEGLEERTKEKGLIIRGWAPQVLILDHESV 354

Qy    356 GAFVTHCGWNSTLEGICAGVPMVTWPVFAEQFFNEKFVTEVLGTGVSVGNNKWLRAASEG 415
      ||||| ||||| : ||||| ||||| ||||| ||| ||| :| :| :| |||
Db    355 GAFVTHCGWNSTLEGVSGGVPMVTWPVFAEQFFNEKLVTEVLKTGAGVGSIQWKRSASEG 414
```

Qy 416 VSREAVTNAVQRVMVGENASEMRKRAKYYKEMARRAVEEGGSSYNGLNEMIEDLSVY 472
| | | | : | : : | | | | | | | | : : | | | | | | : : | | | |
Db 415 VKREAIAKAIKRVMVSEEAEGRNRKAYKEMARKAIEGGGSSYTGLTTLLEDISTY 471

09SXF2

Query Match 69.5%; Score 1749.5; DB 10; Length 476;
Best Local Similarity 68.7%; Pred. No. 1.1e-124;
Matches 332; Conservative 63; Mismatches 79; Indels 9; Gaps 4;

Qy	1	MGKLHIALFPVMAHGHMIPMLDMAKLFTSRGIQTTIISTLAFADPINKARDSGLDIGLSI	60
		: : : : : : : :	
Db	1	MGQLHIVLVPMIAHGHMIPMLDMAKLFSRQVKTIIATPAFAEPIRKARESGHDIGLTT	60
Qy	61	LKFPPEGSGIPDHMVSLDLVTEDWLPKFVESLVLLQEPVEKLI EELKLDCLVSDMFLPWT	120
		: : : : : : :	
Db	61	TKFPPKGSSLPDNIRSLDQVTDLLPHFFRALELLQEPVEEIMEDLKPDCLVSDMFLPWT	120
Qy	121	VDCAAKFGIPRLVFHGTSNFALCASEQMKLHKPYKNVTSdTETfVIPDFPHELKfVVRTQV	180
		: : : : : : :	
Db	121	TDSAAKFGIPRLLFHGTSLFARCFABQMSIQKPYKNVSSDSEPfVLRGLPHEVsfVVRTQI	180
Qy	181	APFQLAE-TENGFSKLMKQMTESVGRSYGVVNSFYELESTYVDYYREVLGRKSWNIGPL	239
		: : : : : : : : : : :	
Db	181	PDYELQEGGDDAFSKMAKQMRDADKKSyGDVINSFEELESEYADYNKNVFGKKAWHIGPL	240
Qy	240	LLSNNNGNEEK-VQRGKESAIGEHECLAWLNSKKQNSVVYVCFGSMATFTPAQLRETAIGL	298
		: : :	
Db	241	KLFNNRAEQKSSQRGKESAI DDHECLAWLNSKKPNSVVYMCFGSMATFTPAQLHETAVGL	300
Qy	299	EESGQEFIWVVKAKNEEEGKGKEEWLPENFEERVKDRGLIIRGWAPQLLILDHPAVGAF	358
		: : : : : : :	
Db	301	ESSGODFIWVVR-----NGGENEDWLPOGFEEIRIKGKGLMIRGWAPQVMILDHPSTGAF	354

Qy 359 VTHCGWNSTLEGICAGVPMVTWPVFAEQFFNEKFVTEVLGTGVS VGNKKWLRAASEGVSR 418
 |||:|||||:||||| ||| |||
 Db 355 VTHCGWNSTLEGICAGLPMVTWPVFAEQFYNEKLVTEVLKTGVS VGNKKWQR-VGEGVGS 413

Qy 419 EAVTNAVQVRVMVGENASEMRKRAKYYKEMARRAVEEGSSYNGLNEMIEDLSVYRAPEKQ 478
 ||| ||:||||: ||| || |||:||||| ||: ||: || ||
 Db 414 EAVKEAVERVMVGDGAAEMRSRALYYKEMARKAVEEGSSYNNLNALIEELSAYVPPMKQ 473

Qy 479 DLN 481
 ||
 Db 474 GLN 476

RESULT 2

P93365

ID P93365 PRELIMINARY; PRT; 476 AA.
 AC P93365;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE IMMEDIATE-EARLY SALICYLATE-INDUCED GLUCOSYLTRANSFERASE.
 GN IS5A.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRIGHT YELLOW 2;
 RX MEDLINE=97000918; PubMed=8843948;
 RA Horvath D.M., Chua N.H.;
 RT "Identification of an immediate-early salicylic acid-inducible tobacco
 RT gene and characterization of induction by other compounds.";
 RL Plant Mol. Biol. 31:1061-1072(1996).
 DR EMBL; U32644; AAB36653.1; -.
 DR Mendel; 9420; Nicta;2542;9420.
 DR InterPro; IPR002213; -.
 DR Pfam; PF00201; UDPGT; 1.
 DR PROSITE; PS00375; UDPGT; 1.
 KW Transferase.
 SQ SEQUENCE 476 AA; 53614 MW; 7C8FD61CEA853F67 CRC64;

Query Match 61.9%; Score 1558.5; DB 10; Length 476;
 Best Local Similarity 59.7%; Pred. No. 3.6e-110;
 Matches 285; Conservative 84; Mismatches 97; Indels 11; Gaps 3;

Qy 1 MGKLHIALFPVMAHGHIIPMLDMAKLFTSRGIQTTIIST---LAFADPINKARDSGLDI 56
 ||:||| ||| ||| ||| |||: |||: | : : |::|
 Db 1 MGQLHIFFFPVMAHGHIPTLDMAKLFASRGVKATIITPLNEFVFSKAIQRNKHGLGIEI 60

Qy 57 GLSILKFPPEGSGIPDHMVSLDLV-TEDWLPKFVESLVLLQEPVEKLIEELKLDCLVSDM 115
 : ::||| :||: || : :: || : :: |||:|:| || : |||:|
 Db 61 EIRLIKFPVAVENGLPEECERLDQIPSDEKLPNFFKAVAMMQEPLEQLIEECRPDCLISDM 120

Qy 116 FLPWTVDCAAKFGIPRLVFHGTSNFALCASEQMKLHKPYKNVTSDETTFVIPDFPHELKF 175
 ||||| | |||| |::||| | ||| :::||:|::||:|::|||:| | ||:|
 Db 121 FLPWTTDTAAKFNIPRIVFHGTSTFFALCVENSVRLNKPKNVSSDSETFVVPDLPHEIKL 180

Qy 176 VRTQVAPFQLAETENGFSKLMKQMTESVGRSYGVVNSFYELESTYVDYYREVLGRKSWN 235
 ||||:|: : | :::| : || :||| | |||||: ||:| :|||:|
 Db 181 TRTQVSPFERSGEETAMTRMIKTVRESDSKSYGVVFNFSFYELETDYVEHYTKVLGRRAWA 240

Qy 236 IGPLLLSNNGNEEKVQRGKESAIGEHECLAWLNSKKQNSVVYVCFGSMATFTPAQLRETA 295
 |||| : | :| :||:|:| :||| |::|| :|||:| |||:| || :|| |
 Db 241 IGPLSMCNRDIEDKAERGKKSSIDKHECLKWLDSKKPSSVVYICFGSVANFTASQLHELA 300

Qy 296 IGLEESGQEFIWVVKKAKNEEEGKGKEEWLPENFEERVKDRGLIIRGWAPQLLILDHPAV 355
 :|:| ||||| |||: : :||| ||| :|:| ||||| |||:| ||| :|
 Db 301 MGVEASGQEFIWVVRTELD-----NEDWLPEGFEERTKEKGLIIRGWAPQVLILDHESV 354

Qy 356 GAFVTHCGWNSTLEGICAGVPMVTWPVFAEQFFNEKFVTEVLGTGVSNGKWLRAASEG 415
 ||||| ||||| |||: ||||| ||||| ||||| || ||:| :|||
 Db 355 GAFVTHCGWNSTLEGVSGGVPMVTWPVFAEQFFNEKLVTEVLKTGAGVGSIQWKRSASEG 414

Qy 416 VSREAVTNAVQRMVGENASEMRKRAKYYKEMARRAVEEGGSSYNGLNEMIEDLSVY 472
 | |||: |:| ||| | | ||| |||||:|:| ||||| || :|:|:| |
 Db 415 VKREAIKAIKRMVSEADGFRNRAKAYKEMARKAIEEGGSSYTGLTTLLEDISTY 471

RESULT 3

P93364

ID P93364 PRELIMINARY; PRT; 476 AA.
 AC P93364;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE IMMEDIATE-EARLY SALICYLATE-INDUCED GLUCOSYLTRANSFERASE.
 GN IS10A.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRIGHT YELLOW 2;
 RX MEDLINE=97000918; PubMed=8843948;
 RA Horvath D.M., Chua N.H.;
 RT "Identification of an immediate-early salicylic acid-inducible tobacco
 RT gene and characterization of induction by other compounds.";
 RL Plant Mol. Biol. 31:1061-1072(1996).
 DR EMBL; U32643; AAB36652.1; -.
 DR Mendel; 9419; Nicta;2542;9419.
 DR InterPro; IPR002213; -.
 DR Pfam; PF00201; UDPGT; 1.
 DR PROSITE; PS00375; UDPGT; 1.
 KW Transferase.
 SQ SEQUENCE 476 AA; 53454 MW; 761A43837A17A232 CRC64;

Query Match 60.3%; Score 1517.5; DB 10; Length 476;

Best Local Similarity 58.5%; Pred. No. 4.6e-107;
Matches 279; Conservative 85; Mismatches 102; Indels 11; Gaps 3;

```

Qy      1 MGKLIHIALFPVMAHGHMIPMLDMAKLFTSRGIQTTIIST----LAFADPINKARDSGLDI 56
      ||:|| ||||| ||||| ||||| |||:: |||:| | : | : : |::|
Db      1 MGQLHFFFPVMAHGHMIPTLDMAKLVASRGVKATIITPLNESVFSKSIQRNKHGLGIEI 60

Qy     57 GLSILKFPPEGSGIPDHMVSLDLV-TEDWLPKFVESLVLLQEPVEKLIEELKLDCLVSDM 115
      : ::||| :|::| |||: ::| || | ::| ::|||:|:| ||| : :||| |||
Db     61 EIRLIKFPVENGLPEECERLDLIPSDDKLPNFFKAVAMMQEPLEQLIEECRPNCCLVSDM 120

Qy    116 FLPWTVDCAAKFGIPRLVFHGTSNFALCASEQMKLHKPYKNVTSDETTFVIPDFPHELKF 175
      ||||| | |||| :||:||||| |||| :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db    121 FLPWTTDTAAKFNMPRIVFHGTSTFFALCVENSIRLNKPFKNVSSDSETFVVPNLPHEIKL 180

Qy    176 VRTQVAPFQLAETENGFSKLMKQMTESVGRSYGVVNSFYELESTYVDYYREVLGRKSWN 235
      |||::|||: : | : ::| : || : |||: | || || | ||::| : |||:|:|
Db    181 TRTQLSPFEQSGETTMTMRIKSVRESDSKSYGVIFNSFNELEHDYVEHYTKVLGRRAWA 240

Qy    236 IGPLLLSNNGNEEKVQRGKESAIGEHECLAWLNSKKQNSVVYVCFGSMATFTPAQLRETA 295
      |||| : | | :| :|||:|:| :||| | :||| : ||||| |||:| | | :| | |
Db    241 IGPLSMCNRDIEDKAERGKQSSIDKHECLKWLDSKKPSSVVYVCFGSVANFTASQLHELA 300

Qy    296 IGLEESGQEFIWVVKKAKNEEEGKGKEEWLPENFEERVKDRGLIIRGWAPQLLILDHPAV 355
      :|:| ||||| |||: : | : ||| | || | :|:| ||||| |||:| ||| :|
Db    301 MGIEASGQEFIWVVRTELD-----NEDWLPEGLEERTKEKGLIIRGWAPQVLILDHESV 354

Qy    356 GAFVTHCGWNSTLEGICAGVPMVTWPVFAEQFFNEKFVTEVLGTGVSVGNKKWLRAASEG 415
      ||||| ||||| |||: | ||||| ||||| ||||| || | | | :| :| | :|||
Db    355 GAFVTHCGWNSTLEGVSGGVPMVTWPVFAEQFFNEKLVTEVLKTGAGVGSIQWKRSASEG 414

Qy    416 VSREAVTNAVQRVMVGENASEMRKRAKYYKEMARRAVEEGGSSYNGLNEMIEDLSVY 472
      | |||: | :||| | | | ||| |||||:|:| |||| | | :||:| |
Db    415 VKREAIKAIKRVMVSEEAEGFRNRAKAYKEMARKAIEGGGSSYTGLTTLLEDISTY 471

```

RESULT 4

Q43526

ID Q43526 PRELIMINARY; PRT; 466 AA.

AC Q43526;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE TWI1 (FRAGMENT).

GN TWI1.

OS Lycopersicon esculentum (Tomato).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;

OC Solanales; Solanaceae; Solanum.

OX NCBI_TaxID=4081;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. MONEYMAKER; TISSUE=WOUNDED LEAF;

RA Truesdale M.R., Doherty H.M., Loake G.J., Mcpherson M.J.,

RA Roberts M.R., Bowles D.J.;

RL Plant Physiol. 112:446-446 (1996)

DR EMBL; X85138; CAA59450.1; -.

LOCUS (LOC): AF117267 GenBank (R)
 GenBank ACC. NO. (GBN): AF117267
 CAS REGISTRY NO. (RN): 225568-86-5
 SEQUENCE LENGTH (SQL): 1819
 MOLECULE TYPE (CI): mRNA; linear
 DIVISION CODE (CI): Plants, fungi, algae
 DATE (DATE): 20 Apr 1999
 DEFINITION (DEF): Malus domestica UDP glucose:flavonoid 3-O-glucosyl transferase (**UFGT1**) mRNA, complete cds.
 SOURCE: apple tree.
 ORGANISM (ORGN): Malus x domestica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
 Tracheophyta; Spermatophyta; Magnoliophyta;
 eudicotyledons; core eudicots; Rosidae; eurosids I;
 Rosales; Rosaceae; Malus
 NUCLEIC ACID COUNT (NA): 488 a 446 c 455 g 430 t
 REFERENCE: 1 (bases 1 to 1819)
 AUTHOR (AU): Lee, J.-R.; Hong, S.-T.; Yoo, Y.G.; Kim, S.-R.
 TITLE (TI): Molecular cloning and expression of anthocyanin biosynthesis genes from 'Fuji apple'
 JOURNAL (SO): Unpublished
 REFERENCE: 2 (bases 1 to 1819)
 AUTHOR (AU): Lee, J.-R.; Hong, S.-T.; Yoo, Y.G.; Kim, S.-R.
 TITLE (TI): Direct Submission
 JOURNAL (SO): Submitted (31-DEC-1998) Life Science, Sogang University, 1 Shinsoo-Dong, Mapo-Gu, Seoul 121-742, South Korea

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..1819	/organism="Malus x domestica" /cultivar="Fuji" /db-xref="taxon:3750" /tissue-type="peel" /note="Malus domestica Borkh"
gene	1..1819	/gene="UFGT1"
CDS	72..1523	/gene="UFGT1" /codon-start=1 /product="UDP glucose:flavonoid 3-O-glucosyl transferase" /protein-id="AAD26203.1" /db-xref="GI:4588779" /translation="MAAPLPiEIEPSSTNGQPHL ADAYNRHVAVVAFPFTHASALLE TVRRLATALPNTLFSFFSTSKSNSSLFSNNSIDN MPRNIRVYDVADGVPEGYVFGKP QEDIELFMNAAPENIRSLDASVADIGKQISCLI TDAFLWFGVHLADELGVPWVTFWI SGLKSLSVHVHTDLIRDTIGTQGITGRENDLIVD KVNVIQGLSNVRIKDLAEGVIFGN LDSVISGMLLQMGRLLPRATAVFMNGFEELELPI PNDLKSKVNKLLNVGPSNVASPLP PLPPSDACLSWLDKQAPSSVVYISFGTVASPAE KEQMAIAEALEATGAPFLWSIKDS CKTPLLNEFLTCTLKSLNGMVVPWAPQPHVLAHD SVGAFVSHCGWNSIMETIAGRVP

ICRPYFADQRLNARMVEEVFEIGVTVEDGVFTRE
GLVKSLEVVLSPESGRKFRDNIKR
VKQLAVEAVGPQGSSTRNFKSLLDIVSGSNIQV"

SEQUENCE (SEQ) :

```
1 gctaactcca ttattccatc agtactgcta cttctattca actccttttc taattagcct
61 tgtaagctgt aatggcagcg ccgctgcccc tcgaaatcga accatcatca actaatggtc
121 aaccccatct cgccgacgcc tacaaccgtc acgtggctgt cgtagccttc cttttacta
181 gccatgcaag cgccttgctt gagaccgtgc gccgcctagc caccgccctt ccaaacactc
241 tcttctcgtt cttcagcact tcaaaatcca acagctctct cttttccaac aacagcattg
301 ataacatgcc gcgtaacata aggggtgtacg atgtggctga cgggggtgccg gaggggtacg
361 ttttcgtggg caagccgcag gaggacatag agctcttcat gaatgccgca ccggaaaaca
421 tccggaggag cttagacgct tccgtggcgg acatcgggaa gcagatcagc tgcttgatca
481 ccgacgcctt cctttggttt ggagtccact tggctgacga gttgggagtg ccttgggtca
541 ctttctggat ctccggactc aaatccctct ccgttcatgt gcatactgat ctcatccgcg
601 acactattgg aactcaaggc attacaggtc gtgaaaacga cctcatcgtc gacaaaaatg
661 ttaacatcca aggtctctcc aatgtacgaa tcaaagactt agcgggaagga gtcattttcg
721 gaaacttgga ctcggttaatt tccggcatgc tacttcagat gggacggctg ctcccccg
781 ccaccgcagt tttcatgaac ggcttcgaag aattggaact ccccatacca aacgacctaa
841 agtccaaagt caacaaactc ctcaacgtag gaccttccaa cgtagcatcc ccgctgccac
901 cgctgccgcc atcagatgct tgcttgctcat ggctagacaa gcaacaggct ccacctccg
961 tcgtgtacat aagcttcggg acagtggcga gccagcgga gaaggagcag atggcaatag
1021 cggaggccct ggaagccacc ggagcaccct tcttgtggtc tatcaaggac agctgcaaga
1081 caccgttget gaacgagttc ttgacaaaaa cattgtcaaa gctgaacggg atgggtggtc
1141 cgtgggctcc acagccgcac gtactggccc acgattcggg cggagccttc gtgtcgcat
1201 gcggttgaa ctcgataatg gagactatag caggacgggt gccatgatt ttagggccat
1261 attttgcaga ccagaggctt aatgcaagga tgggtggagg ggtgtttgag atcggggtaa
1321 ccgtggagga tggagttttt accagggagg ggctggtaaa aagcttgaa gtgttttgt
1381 cgctgaaag tgggaggaaa ttcagagaca atataaagag ggtcaaaca ctggcagtag
1441 aggcggttg accacaaggg agctccactc ggaacttcaa atcgctgttg gacatcgat
1501 caggatccaa ttatcaagta tagtacgggg accataaata tgtagcacta aaaatacaca
1561 ccagctgcaa tagctgttcg ttgctcaata ataattgtagc acgctacgat ctaccctacg
1621 agcggacata tttaggcgcg gtttggtatt gatgtgcttt gaaaaaaga tgctgtgaga
1681 ataagcggt gtgctgtgaa aataagcggc tgtgaaataa atcagcagag tgtttgataa
1741 acgtttttgt aaaagtgtt ttggaaagaa aaaaaacagt ctaatagtgt gtcttttcat
1801 taaaaaaaaa aaaaaaaaaa
```